Screening for Prostate Cancer: An Update of the Evidence for the U.S. Preventive Services Task Force

Russell Harris, MD, MPH, and Kathleen N. Lohr, PhD

Background: In U.S. men, prostate cancer is the most common noncutaneous cancer and the second leading cause of cancer death. Screening for prostate cancer is controversial.

Purpose: To examine for the U.S. Preventive Services Task Force the evidence of benefits and harms of screening and earlier treatment.

Data Sources: MEDLINE and the Cochrane Library, experts, and bibliographies of reviews.

Study Selection: Researchers developed eight questions representing a logical chain between screening and reduced mortality, along with eligibility criteria for admissible evidence for each question. Admissible evidence was obtained by searching the data sources.

Data Extraction: Two reviewers abstracted relevant information using standardized abstraction forms and graded article quality according to Task Force criteria.

Data Synthesis: No conclusive direct evidence shows that screening reduces prostate cancer mortality. Some screening tests can detect prostate cancer at an earlier stage than clinical detection. One study provides good evidence that radical prostatectomy reduces disease-specific mortality for men with localized prostate cancer detected clinically. No study has examined the additional benefit of earlier treatment after detection by screening. Men with a life expectancy of fewer than 10 years are unlikely to benefit from screening even under favorable assumptions. Each treatment is associated with several well-documented potential harms.

Conclusions: Although potential harms of screening for prostate cancer can be established, the presence or magnitude of potential benefits cannot. Therefore, the net benefit of screening cannot be determined.

For author affiliations, see end of text.
See related article on pp 915-916.

The American Cancer Society estimates that 189,000 men will receive a diagnosis of prostate cancer in 2002 and that 30,200 men will die of the disease (1). Many more men receive a diagnosis of prostate cancer than die of it (lifetime risk, about 1 in 6 vs. about 1 in 29). Among types of cancer, only lung cancer kills more men each year. The cause of prostate cancer is unknown, and the best-documented risk factors (age, ethnicity, and family history) are not modifiable. The burden of prostate cancer falls disproportionately on men who are older or black. The median age at diagnosis is approximately 71 years, and the median age at death is 78 years (2). More than 75% of all cases of prostate cancer are diagnosed in men older than 65 years of age, and 90% of deaths occur in this age group (2, 3). Incidence is approximately 60% higher and mortality rate is twofold higher in black men than in white men (2). Asian-American men and Hispanic men have lower incidence rates than non-Hispanic white persons (3).

Although approaches to primary prevention of prostate cancer are being tested, to date none are known to be effective. The most common strategy for reducing the burden of prostate cancer is screening, but screening remains controversial. Many studies on this topic have been published since 1996, when the U.S. Preventive Services Task Force (USPSTF) last examined prostate screening (4). To assist the USPSTF in updating its recommendation, the Research Triangle Institute–University of North Carolina Evidence-based Practice Center performed a systematic review of the evidence on screening for prostate cancer.

Methods

Using USPSTF methods (5), we developed an analytic framework and eight key questions to guide our literature search. Because we found no direct evidence connecting screening and reduced mortality, we searched for indirect evidence on the yield of screening, the efficacy and harms of various forms of treatment for early prostate cancer, and the costs and cost-effectiveness of screening. We developed eligibility criteria for selecting relevant evidence to answer the key questions (Table 1). We examined the critical literature from the 1996 USPSTF review and used search terms consistent with the eligibility criteria to search the MEDLINE database and Cochrane Library for English-language reviews and relevant studies published between 1 January 1994 and 15 September 2002.

The first author and at least one trained assistant reviewed abstracts and articles to find those that met the eligibility criteria. For these studies, the two reviewers abstracted relevant information using standardized abstraction forms. We graded the quality of all included articles according to USPSTF criteria (5). The authors worked closely with two members of the USPSTF throughout the review and periodically presented reports to the full USPSTF. We distributed a draft of the systematic evidence review to experts in the field and relevant professional organizations and federal agencies for broad-based external peer review and made revisions based on the feedback. We then revised the full systematic evidence review into this manuscript. A more complete account of the methods of this review can be found in the Appendix (available at

This evidence report was funded through a contract to the Research Triangle Institute–University of North Carolina Evidence-based Practice Center from the Agency for Healthcare Research and Quality. Staff of the funding agency and members of the USPSTF contributed to the study design, reviewed draft and final manuscripts, and made editing suggestions.

RESULTS

Direct Evidence That Screening Reduces Mortality
Randomized, Controlled Trials

Labrie and colleagues (7) completed the first randomized, controlled trial (RCT) of prostate cancer screening with more than 46,000 men. At the end of 8 years of follow-up, approximately 23% of the invited group and 6.5% of the not-invited group had been screened with prostate-specific antigen (PSA) testing and digital rectal examination (DRE). Prostate cancer death rates did not differ between groups (4.6 vs. 4.8 deaths per 1000 persons, respectively).

Two other RCTs of prostate cancer screening, both initiated in 1994, are ongoing: the U.S. National Cancer Institute Prostate, Lung, Colorectal, and Ovary Trial and the European Randomized Study of Screening for Prostate Cancer. Neither study will have data on mortality for several more years.

Case–Control Studies

Three well-conducted, nested case-control studies (two since 1994) examined the relationship between chart review documentation of DRE and advanced prostate cancer or death from prostate cancer. Two studies found no relationship (8, 9). The third study found that men who died of prostate cancer had fewer DREs in the years before diagnosis (odds ratio indicating a protective effect of DRE, 0.51 [95% CI, 0.31 to 0.84]) (10).

Why results from these otherwise similar studies differ is not clear. The three studies depended on large databases and on individual medical records. They defined cases slightly differently and used different approaches to differentiate screening DRE from diagnostic DRE. Because such studies are complex in design, we were not able to determine whether one method was more accurate than another (11). All three studies were small, and all were consistent

---

**Table 1. Key Questions, Inclusion Criteria, and Articles Meeting Criteria***

<table>
<thead>
<tr>
<th>Key Question</th>
<th>Inclusion Criteria</th>
<th>Articles Meeting Criteria, n</th>
</tr>
</thead>
</table>
| All          | Published from 1 January 1994 to 15 September 2002  
                English language  
                MEDLINE, Cochrane  
                Human participants | 1 RCT  
                        2 case–control  
                        15 ecologic |
| 1. Efficacy of screening (direct evidence) | RCT; case–control study; or ecologic evidence directly connecting screening with health outcomes | 15 |
| 2. Yield of screening                  | Unselected population without prostate cancer  
                Screening test offered to all  
                Work-up offered to all with positive results on screening tests  
                Screening test compared with a valid reference standard | 35 |
| 3. Efficacy of radical prostatectomy   | RCT; clinically localized disease  
                Follow-up ≥2 years  
                ≥75% of patients followed | 1 |
| 4. Efficacy of radiation therapy       | Health outcomes | 0 |
| 5. Efficacy of androgen deprivation therapy | | 2 |
| 6. Efficacy of watchful waiting        | | 1 |
                Use of valid measurement instrument  
                Follow-up from pretreatment to at least 12 months post-treatment or  
                Comparison with similar untreated control group at least 12 months post-treatment | 32 |
| 8. Costs and cost-effectiveness of treatment | Valid assessment of costs of screening and treatment  
                Assess direct and indirect costs  
                Cost-effectiveness, cost-benefit, cost-utility  
                Modeling studies | 2 |

* RCT = randomized, controlled trial.
Changes in prostate cancer treatment during the late 1980s and early 1990s included higher rates of radical prostatectomy, development of luteinizing hormone–releasing hormone (LHRH) agonists (allowing improved androgen deprivation therapy without castration), and refinements in radiation therapy. Such changes may explain the reduction in prostate cancer mortality. A recent study by Bartsch and coworkers (27), for example, documented a greater reduction in prostate cancer mortality in the Austrian state of Tyrol, which had instituted a free PSA screening program, compared with the rest of Austria. This finding could be a consequence of the screening program, changes in treatment that accompanied the screening program, misattribution of cause of death, or some combination of the three.

Accuracy of Screening

Three problems complicate any attempt to determine the accuracy of screening tests for prostate cancer. First, research has yet to clarify which tumors screening should target. Second, the reference standard (prostate biopsy) for diagnosing prostate cancer after positive results on a screening test is imperfect. Third, few studies perform biopsy on men with negative results on screening tests.

Prostate cancer is a heterogeneous tumor. Different cases of prostate cancer have widely varying growth rates and potential for causing death. Ideally, prostate cancer screening would target only tumors that would cause clinically important disease. Currently available prognostic markers can distinguish a small number of men with excellent prognosis for long-term survival and a small number of men with poor prognosis for long-term survival (28). However, they cannot help us correctly categorize the prognosis of those in the middle category, which includes most men with prostate cancer (29–33). Since research has not yet clearly defined the characteristics of clinically important prostate cancer, we do not know what the specific target of screening should be.

The usual reference standard used in prostate cancer screening studies, transrectal needle biopsy of the prostate, is imperfect for two reasons. First, it misses some cases of cancer; 10% to 30% of men who have negative results on an initial series of biopsies have cancer on repeated biopsy series (34–39). Thus, some men categorized as not having cancer actually have it, falsely lowering the test’s measured sensitivity. Second, in clinical practice and research, a “biopsy” is actually four to six (or more) biopsies. Many biopsy specimens are obtained, most from normal-appearing areas of the prostate. An analysis of this practice concluded that up to 25% of apparently PSA-detected tumors and more than 25% of apparently DRE-detected tumors were likely to have been detected by serendipity, that is, as an incidental finding on a blind biopsy (40). Thus, some men who are categorized as having cancer detected by screening actually have serendipity-detected cancer. This error falsely increases sensitivity.
In addition to problems of the accuracy of the reference standard, few studies perform biopsy on men who have negative results on screening tests. This reduces our ability to determine the number of false-negative screening tests and to calculate sensitivity. Most studies use several noninvasive tests together, measuring the sensitivity of one test against the combined findings of all tests. The extent to which these combined tests actually detect all important cancer is unknown. This bias probably leads to an overestimate of sensitivity.

Screening Methods
Prostate-Specific Antigen Testing

An analysis from the Physicians’ Health Study avoided some of the bias of the problematic reference standard by using longitudinal follow-up instead of biopsy (41). In this study, which used a PSA cut-point of 4.0 ng/mL or higher, the sensitivity for detecting cancer appearing within 2 years after screening was 73.2%. Although the study calculated sensitivity separately for aggressive (that is, extracapsular or higher grade) and nonaggressive (that is, intracapsular and lower grade) cancer (sensitivity, 91% vs. 56%), it is not clear that these categories correspond to clinically important and clinically unimportant tumors. Among men who did not receive a diagnosis of prostate cancer in those 2 years, 14.6% had an initial PSA level of 4.0 ng/mL or greater, corresponding to a specificity of 85.4%.

Other studies have provided similar estimates of sensitivity and specificity for PSA level with a cut-point of 4.0 ng/mL (42–44). Specificity for PSA screening is lower among men with larger prostate glands, including the large number of older men with benign prostatic hyperplasia. One study of four carefully chosen samples found that the likelihood ratios for various PSA levels were much lower among men with benign prostatic hyperplasia than among men without benign prostatic hyperplasia (45). Thus, the PSA test is not as accurate in detecting cancer in men with benign prostatic hyperplasia as in those without.

Because of the reduced specificity in older men with benign prostatic hyperplasia, some experts have proposed that the PSA cut-point be adjusted for age, with higher cut-points for older men and lower cut-points for younger men (46). Such a strategy increases sensitivity and lowers specificity in younger men, while the reverse is true in older men. Experts disagree about whether this strategy would improve health outcomes (47, 48).

Some experts have also proposed decreasing the cut-point defining an abnormal PSA level from 4.0 ng/mL to 3.0 ng/mL (or even 2.6 ng/mL) for all men (44, 49, 50). This approach results in more biopsies and more cancer detected (49, 51–55). Because of the uncertainty about the definition of clinical importance, the value of this increased detection is unknown.

In the serum, PSA circulates in two forms: free and complexed with such molecules as α1-antichymotrypsin. Men with prostate cancer tend to have a lower percentage of free PSA than men without prostate cancer (56, 57). In research, the percentage of free PSA (or a similar test measuring the level of complexed PSA) (58–62) has mainly been used to increase the specificity of screening by distinguishing between men with PSA levels of 4.0 ng/mL to 9.9 ng/mL who should undergo biopsy and those who should not (52, 63–70). Different studies have suggested different cut-points for percentage of free PSA; a lower cut-point avoids more biopsies but also misses more cancer. High cut-points (for example, 25%) would avoid about 20% of biopsies, and the probability of cancer at that cut-point is about 8% (66). In practice, it is not clear whether this probability would be low enough for men and their physicians to forgo biopsy (71).

Men with prostate cancer have a greater increase in PSA level over time than men without cancer (72). It is unclear, however, whether examining the annual rate of change in PSA level (PSA velocity) improves health outcomes or reduces unnecessary biopsies (47, 73). Because of intraindividual variation, PSA velocity is useful only in men who have three or more tests of PSA level over a period of 1 to 3 years (47, 74, 75).

Digital Rectal Examination

It is more difficult to detect cancer with DRE than with PSA. A meta-analysis examining studies of DRE in unselected samples screened by both PSA and DRE found a sensitivity of 59% (64%) for the four best studies (76). Digital rectal examination detects cancer in some men with PSA levels below 4.0 ng/mL (positive predictive value, about 10% according to one large study [63]) or even 3.0 ng/mL, but the tumors were usually small and well differentiated (77). Digital rectal examination has limited reproducibility (78).

Yield of Large Screening Programs

Using six studies of screening with a single PSA test or with PSA and DRE among large, previously unscreened samples, we were able to estimate the yield of a new screening program among men in different age groups (79) who had not previously been screened (7, 44, 49, 63, 64, 79–83). The Figure gives estimates of positive test results and cases of cancer detected after screening with PSA alone or screening with PSA and DRE among men in their 60s. Men in their 50s have fewer positive test results and cases of cancer detected, while men in their 70s have more.

The percentage of participants with a PSA level of 4.0 ng/mL or higher ranged from about 4% (79) among men in their 50s to about 27% (64) among men in their 70s. The percentage of men who had a PSA level of 4.0 ng/mL or higher or abnormal results on DRE ranged from 15% among younger men to 40% among older men (79). Few other screening tests have such a high percentage of positive results (63, 64, 81, 82).

In the screening studies, some men with abnormal results on screening tests did not undergo biopsy. If we assume that biopsy is performed on all men with an abnor-
mal result on a screening test and that the rate of cancer
detection is the same as for men who undergo study bi-
opsy, we estimate that the percentage of all men screened
who would have prostate cancer detected would range
from approximately 1.5% (PSA screening alone for men in
their 50s) (81) to 10% (PSA and DRE screening for men
in their 70s) (44).

The number of cancer cases detected would be smaller after repeated annual screening. * For men in their 50s, n = 50; for men in their 70s, n = 270.
† For men in their 50s, n = 150; for men in their 70s, n = 400. ‡ For men in their 50s, n = 17; for men in their 70s, n = 90. § For men in their 50s, n = 30; for men in their 70s, n = 100. ¶ For men in their 50s, n = 12; for men in their 70s, n = 63. In all footnotes, numbers are approximate.

Yield with Different Screening Intervals

Rates of positive results and cancer detection decrease
on screening a year after the initial screening round (7, 49,
80, 83–85). In one study, approximately 26% of men with
a PSA level of 4.0 ng/mL or greater had prostate cancer
after the first round of screening and approximately 6.2%
had cancer after subsequent rounds (7). Other studies have
concluded that annual screening confers little gain com-
pared with intervals of at least 2 years (86), especially for
the 70% of the population with PSA levels of 2.0 ng/mL or
less (41, 87).

Effectiveness of Current Treatments for Localized
Disease

Radical Prostatectomy

Since 1991, radical prostatectomy has been the most
common treatment for clinically localized prostate cancer.
It is the initial treatment for more than one third of pa-
tients with new diagnoses, most commonly men 75 years

www.annals.org
of age or younger (2). The procedure is usually performed with curative intent in men who have a life expectancy of at least 10 years.

One well-conducted RCT compared radical prostatectomy with “watchful waiting” (in which treatment is not given initially but is reserved for progressive or symptomatic disease) among men with clinically detected prostate cancer (88). About 75% of the men in this study had palpable cancerous tumors, few of which were the size usually detected by PSA screening (73, 84, 89). After 8 years of follow-up, 7.1% of the radical prostatectomy group had died of prostate cancer compared with 13.6% of men in the watchful waiting group (relative hazard, 0.50 [CI, 0.27 to 0.91]). The absolute difference in prostate cancer mortality was 6.6% (CI, 2.1% to 11.1%) (number needed to treat for benefit, 17). The groups did not differ in all-cause mortality.

No other well-conducted RCT has compared any other treatment with radical prostatectomy for clinically localized prostate cancer. One ongoing RCT, the prostatectomy Intervention versus Observation Trial in the United States, will publish results in the future (90, 91).

One observational study that used internal controls and data from the Surveillance, Epidemiology, and End Results program provided information on the effectiveness of radical prostatectomy relative to other treatments (92). For men with well-differentiated cancer, 10-year disease-specific survival did not differ between the radical prostatectomy group and the age-matched radiation or watchful waiting groups. Disease-specific survival was slightly higher for the radical prostatectomy group in men with moderately differentiated tumors (radical prostatectomy group, 87%; radiation group, 76%; watchful waiting group, 77%) and was much higher for men with poorly differentiated cancer (radical prostatectomy group, 67%; radiation group, 53%; watchful waiting group, 45%). Other cohort studies without controls have found similar survival rates after radical prostatectomy (33, 93–97).

**Radiation Therapy**

Radiation therapy is the second most commonly used treatment for nonmetastatic prostate cancer and is the most common treatment for men 70 to 80 years of age (2). The two common types of radiation therapy reviewed here are external-beam radiation therapy (EBRT) and brachytherapy, the insertion of radioactive pellets directly into prostate tissue.

No well-conducted RCT with clinical outcomes compares EBRT with any other therapy for clinically localized prostate cancer. In the large cohort study discussed earlier (92), 10-year disease-specific survival rates in the EBRT group were similar to those in the watchful waiting group for men with well-differentiated and moderately differentiated tumors but were higher for men with poorly differentiated cancer (92).

Brachytherapy is most often used alone for men with well or moderately differentiated intracapsular prostate cancer or in combination with EBRT for men with more aggressive cancer. No RCT with clinical outcomes compared brachytherapy with any other treatment for prostate cancer. Two observational studies involving 100 or more patients with clinically localized prostate cancer found high survival rates for patients treated with radioactive gold or iodine seeds (98, 99).

**Androgen Deprivation Therapy**

The traditional approach to androgen deprivation therapy has been surgical bilateral orchietomy. A newer approach uses LHRH agonists (for example, goserelin or leuprolide), a group of drugs that stimulate the release of luteinizing hormone from the pituitary gland. Paradoxically, when used clinically, LHRH agonists result in down-regulation of pituitary receptors, thus markedly reducing the level of testosterone production to that of a castrated man. Luteinizing hormone–releasing hormone agonists have been used clinically since the late 1980s.

Two well-conducted RCTs compared clinical outcomes between men with clinically localized prostate cancer who were treated with androgen deprivation therapy (with orchiectomy [100] or estramustine [101]) and men treated with EBRT. Androgen deprivation therapy either increased overall survival (100) or reduced clinical recurrence (101); outcomes improved primarily among men who had lymph node involvement.

Four additional RCTs of androgen deprivation therapy (with LHRH agonists) as an adjuvant to EBRT or radical prostatectomy for locally advanced prostate cancer found statistically significant improved overall survival (10% to 20% absolute difference) in men who received androgen deprivation therapy (102–108). Another RCT of immediate versus deferred androgen deprivation therapy (with orchiectomy or LHRH agonists) and no other treatment found improved survival (8% absolute difference) for the immediate therapy group in men who had a new diagnosis of locally advanced prostate cancer (109).

**Watchful Waiting**

The term watchful waiting implies that no treatment is given initially but that the patient is followed for evidence of progressive or symptomatic disease, for which treatment might be offered. Because the only well-conducted RCT that compares watchful waiting and more aggressive treatment examined men with prostate cancer detected clinically rather than by screening (88), the best information about the outcomes of watchful waiting comes from observational studies of men who, for various reasons, were not treated for prostate cancer. These studies also provide information about the natural history of the disease.

Four well-conducted retrospective cohort studies (29, 30, 110, 111) and one pooled analysis of six other cohort studies (28) provide information about survival with untreated prostate cancer. Men with well-differentiated, clin-
Annals of Internal Medicine
Volume 137
923

Screening for Prostate Cancer: An Update of the Evidence

Table 2. Harms of Treatment

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Men with Reduced Sexual Function</th>
<th>Men with Urinary Problems</th>
<th>Men with Bowel Problems</th>
<th>Men with Other Symptoms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Radical prostatectomy</td>
<td>20–70</td>
<td>15–50</td>
<td>6–25</td>
<td>18†</td>
</tr>
<tr>
<td>External-beam radiation therapy</td>
<td>20–45</td>
<td>2–16</td>
<td>6–12†</td>
<td>18†</td>
</tr>
<tr>
<td>Brachytherapy</td>
<td>36†</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Androgen deprivation therapy (LHRH agonists)</td>
<td>40–70</td>
<td></td>
<td></td>
<td>5–25</td>
</tr>
</tbody>
</table>

* Values are percentages of men treated who had side effects at least 12 months after treatment. LHRH = luteinizing hormone–releasing hormone.
† These findings are less certain than other entries because they are based on less or inferior evidence.

Clinically localized prostate cancer have excellent long-term survival, with little or no reduction in survival compared with similar men without prostate cancer. Men with poorly differentiated cancer have reduced survival. In one study, 10-year survival was 17% in men with poorly differentiated cancer and 47% in age-matched controls without prostate cancer (92).

Because most prostate cancer detected today by screening is moderately differentiated, survival of men with this type of tumor is important to the debate about screening. On the standard histologic grading system for prostate cancer, these men have tumors with Gleason scores of 5 to 7. Gleason scores range from 2 to 10; lower scores indicate well-differentiated patterns, and higher scores indicate more poorly differentiated tumors.

The most detailed analysis of men with untreated, clinically localized, moderately differentiated cancer found that 15-year prostate cancer–specific survival rates ranged from 30% (Gleason score of 7 in men 50 to 59 years of age) to 94% (Gleason score of 5 in men 50 to 59 years of age) (29, 110). Men in their 70s had survival rates similar to those of men in their 50s for tumors with a Gleason score of 5 but much better survival for tumors with a Gleason score of 7 (58% vs. 30%). Because men in this study received their diagnoses in the pre-PSA era, survival would probably be even better in similar men receiving diagnoses today given the “lead time” added by earlier detection.

Harms of Treatment

Because harms of treatment are experienced by the men themselves, and because men may have problems that are similar to treatment harms but are not attributable to treatment, we prioritized evidence that measured patients’ perceptions of their own function. For comparison, we used an untreated group or the same men examined before and at least 12 months after treatment.

Radical Prostatectomy

Thirty-day mortality rates after radical prostatectomy range from 0.3% to 1% for most men and may be higher for men older than 80 years of age (112–116). The primary long-term adverse effects of radical prostatectomy include erectile dysfunction and urinary incontinence (Table 2). At least 20%, perhaps as many as 70%, of men have worsened sexual function as a result of radical prostatectomy (114, 117–135). Fifteen percent to 50% of men who had a radical prostatectomy had some urinary problems 1 year later (112, 114, 117, 119, 121, 123, 126, 127, 131, 132, 135–138). Current evidence is mixed about the extent to which, outside of excellent academic centers (95, 124), the newer nerve-sparing procedure reduces complication rates.

Radiation Therapy

Twenty percent to 45% of men with no erectile dysfunction and 2% to 16% of men with no urinary incontinence before EBRT developed dysfunction 12 to 24 months afterward (Table 2) (117, 123, 126–131, 135, 137, 139–146). Six percent to 25% of men who had no bowel dysfunction before EBRT reported marked problems 12 or more months afterward (117, 123, 125, 126, 130, 131, 135, 137, 140, 142, 144–147). The evidence is mixed about whether newer techniques, including three-dimensional conformal EBRT, reduce the frequency of urinary or bowel side effects.

Compared with EBRT or radical prostatectomy, fewer high-quality studies of the harms of brachytherapy have been completed. Our estimates are therefore less precise for this treatment. Among men who were potent before treatment, about 21% are impotent and 36% have decreased erectile function 3 years after brachytherapy (148, 149). A majority of men will have distressing urinary symptoms in the first months after brachytherapy, and 6% to 12% will have such symptoms 1 year later. Up to 25% of men will have some lack of urinary control 12 months after brachytherapy (120, 149–151). Approximately 18% of men will have diarrhea 1 year later (120), and 19% will have some persistent rectal bleeding 12 to 28 months later (152).

Androgen Deprivation Therapy

We focused on the harms of LHRH agonists because the effectiveness studies we reviewed primarily used this...
type of androgen deprivation therapy. No study has examined reports from the same patients beginning before androgen deprivation therapy and extending for at least 1 year. Our best information comes from two large national studies (153–155) and a systematic review (156, 157). Compared with untreated men, 40% to 70% of men who were sexually active before treatment were not sexually active afterward (Table 2). Five percent to 25% of men had breast swelling, and 50% to 60% had hot flashes. Mean scores on quality-of-life indices are lower for men treated with androgen deprivation therapy (154). One RCT of LHRH adjuvant therapy found similar results (107). Potential long-term complications of LHRH therapy include lack of vitality, anemia, and osteoporosis (155, 158, 159). The frequency and severity of these complications are not yet clear.

Quality of Life

Litwin and colleagues (129) compared overall quality-of-life scores among controls and men with prostate cancer within treatment groups. Although they found the same differences in specific symptoms as noted earlier, they found no differences among groups (either among treatment groups or between men with and without prostate cancer) in overall quality of life.

Cost-Effectiveness of Screening

Given the uncertainties about the existence and magnitude of benefits, the cost-effectiveness of screening for prostate cancer has been difficult to calculate. A 1993 decision analysis, which made optimistic assumptions about benefit from screening and early treatment, found little or no benefit for men with well-differentiated tumors (160). For men with moderately or poorly differentiated cancer, screening and early treatment could offer as much as 3.5 years of improvement in quality-adjusted life expectancy, again using the most optimistic assumptions. Even with optimistic assumptions, however, men 75 years of age and older were not likely to benefit from screening and aggressive treatment. One major reason is that any benefits of screening are expected to accrue some years in the future, after many men in this age group have died of some other condition. Two subsequent decision analyses have reached the same conclusions (161, 162).

In 1995, Barry and coworkers (163) published a cost-effectiveness analysis using favorable screening assumptions. The marginal cost-effectiveness of screening men 65 years of age with PSA and DRE, without adjustment for quality of life and without discounting benefits, was between $12500 and $15000 per life-year saved. Changing only a few assumptions, however, quickly increased the marginal cost-effectiveness ratio to above $100000 per life-year saved. This ratio would be even less favorable if a decrement in quality of life associated with the harms of treatment were considered. In 1997, these investigators updated their model with newer data and further assumptions favorable to screening; findings were similar (164).

DISCUSSION

Prostate-specific antigen testing and, to a lesser extent, DRE can detect prostate cancer at an earlier stage than it could be detected clinically. A major problem in considering the utility of screening, however, is the heterogeneity of prostate cancer itself. The large discrepancy between prostate cancer diagnoses and deaths indicates that some and probably most tumors detected by screening are clinically unimportant. Because precise evidence regarding the prognosis of prostate cancer of various types is lacking, researchers have not been able to define the most appropriate targets of screening, that is, the types of cancer that will cause clinical symptoms and death and that can be treated better if detected earlier.

The efficacy of various types of treatment for clinically localized prostate cancer, and especially for the types of localized prostate cancer detected by screening, is largely unknown. Although one RCT found that radical prostatectomy reduced prostate cancer mortality compared with watchful waiting among men with symptomatic localized cancer, the magnitude of any additional benefit of detection and earlier treatment due to screening is still unknown. We lack direct evidence that EBRT, brachytherapy, or androgen deprivation therapy is effective for clinically localized cancer. Each treatment for prostate cancer is associated with various potential harms, including sexual, urinary, and bowel dysfunction.

The costs of a screening program for prostate cancer are potentially high. If treatment is extremely efficacious, then the cost-effectiveness of screening men 50 to 69 years of age may be reasonable; if treatment is less efficacious, the results may be net harm and high costs. Assuming that any potential benefit to screening accrues only after some years, men with a life expectancy of fewer than 10 years are unlikely to benefit. Because prostate cancer incidence and mortality rates are higher among black men, beneficial screening could have a larger absolute benefit in this ethnic group than in white men. The same uncertainties about screening, however, would apply.

Two RCTs of screening are in progress. Because of the problem of screening in control groups, however, some experts fear that even these trials may not provide a definitive answer about screening efficacy. If these trials find a reduction in prostate cancer mortality, further research will be required to determine whether the benefits outweigh the harms and costs for individuals or as a general policy. Research can help by developing new screening and treatment approaches that minimize harms and costs. If the trials show no benefit, research on other approaches to disease control, such as chemoprevention, will be necessary. In the interim, the efficacy of screening for prostate cancer remains uncertain.

From University of North Carolina at Chapel Hill, Chapel Hill, and Research Triangle Institute, Research Triangle Park, North Carolina.
Screening for Prostate Cancer: An Update of the Evidence  

**Clinical Guidelines**

**Disclaimer:** The authors of this article are responsible for its contents, including any clinical or treatment recommendations. No statement in this article should be construed as an official position of the U.S. Agency for Healthcare Research and Quality or the U.S. Department of Health and Human Services.

**Acknowledgments:** The authors acknowledge the continuing support of David Atkins, MD, MPH, and the guidance and feedback of Cynthia D. Mulrow, MD, MPH, University of Texas at San Antonio, Paul S. Frame, MD, Cohocton, New York, and other members of the USPSTF. They thank the staff of the Research Triangle Institute–University of North Carolina Evidence-based Practice Center, including Audrina J. Bunton, BA, Linda Lux, MPA, and Sonya Sutton, BSPH, for substantive and editorial work and Loraine Monroe for secretarial assistance. They also thank Rainer Beck, MD, Paul Godley, MD, MPH, and Kenneth Fink, MD, MPH, for assistance in reviewing articles. In addition, they thank the staff of the University of North Carolina at Chapel Hill and the Cecil G. Sheps Center for Health Services Research, including Timothy S. Carey, MD, MPH, Anne Jackman, MSW, Lynn Whitener, MSLS, DrPH, and Carol Kranov. They also thank the peer reviewers of the complete systematic evidence review: Michael Barry, MD, Medical Practices Evaluation Center, Boston, Massachusetts; John W. Feighnert, MD, MSC, Canadian Task Force on Preventive Health Care, London, Ontario, Canada; Theodore Ganiats, MD, University of California at San Diego, La Jolla, California; Marc B. Garnick, MD, Beth Israel Deaconess Medical Center, Boston, Massachusetts; Richard Hoffman, MD, MPH, Albuquerque Veterans Affairs Medical Center, Albuquerque, New Mexico; Barnett Kramer, MD, National Institutes of Health, Bethesda, Maryland; David Lush, MD, Medical College of Pennsylvania, Philadelphia, Pennsylvania; Curtis Mettlin, MD, Roswell Park Cancer Institute, Buffalo, New York; and Ian M. Thompson Jr., MD, University of Texas Health Sciences Center at San Antonio, San Antonio, Texas.

**Grant Support:** This study was conducted by the Research Triangle Institute–University of North Carolina Evidence-based Practice Center under contract to the Agency for Healthcare Research and Quality, Rockville, Maryland (contract no. 290-97-0011, task order 3).

**Requests for Single Reprints:** Reprints are available from the Agency for Healthcare Research and Quality Web site (www.ahrq.gov/clinic/uspstff.htm) or the Agency for Healthcare Research and Quality Publications Clearinghouse.

Current author addresses, the Appendix, Appendix Table, and Appendix Figures are available at www.annals.org.

**References**


3 December 2002 Annals of Internal Medicine Volume 137 • Number 11 925


44. Catalona WJ, Smith DS, Ornstein DK. Prostate cancer detection in men with serum PSA concentrations of 2.6 to 4.0 ng/ml and benign prostate examination. Enhancement of specificity with free PSA measurements. JAMA. 1997;277:1452-5. [PMID: 9145717]


Carter HB, Epstein JI, Chan DW, Fozard JL, Pearson JD. Recommended prostate-specific antigen testing intervals for the detection of curable prostate cancer. JAMA. 1997;277:1456-60. [PMID: 9145718]


Schmidt JD, Gibbons RP, Murphy GP, Bartolucci A. Evaluation of adjuvant estramustine phosphate, cyclophosphamide, and observation only for node-positive patients following radical prostatectomy and definitive irradiation. Investigators of the National Prostate Cancer Project. Prostate. 1996;28:51-7. [PMID: 8945261]


APPENDIX

This appendix documents procedures that the Research Triangle Institute–University of North Carolina Evidence-based Practice Center (EPC) staff used to develop this report on screening for prostate cancer. During preparation of the evidence report, we collaborated with two current members of the USPSTF who served as liaisons to the EPC topic team. We first document the analytic framework and key questions developed at the beginning of the review. We then describe the inclusion and exclusion criteria for admissible evidence, our strategy for literature search and synthesis, and our approach to developing the final summary of the evidence.

Analytic Framework and Key Questions

The analytic framework (Appendix Figure 1) describes the relationship between screening and treating patients in a clinical setting and reduced morbidity or mortality from prostate cancer. We examined one overarching question (key question 1, linking screening and ultimate health outcomes) and eight additional questions pertaining to specific links in the analytic framework. The key questions were as follows: 1) What are the health outcomes (both type and magnitude) of screening a defined population for prostate cancer compared with not screening? 2) What is the yield of screening for prostate cancer (that is, accuracy and reliability of screening tests, prevalence of undetected cancer in various populations)? 3–6) What are the health outcomes associated with treating clinically localized prostate cancer with radical prostatectomy, external-beam radiation therapy or brachytherapy, androgen deprivation, or watchful waiting? 7) What harms are associated with these treatments of clinically localized prostate cancer? 8) What costs are associated with screening for and early treatment of prostate cancer? Have studies modeled the potential benefits of screening? What is the cost-effectiveness of screening for prostate cancer? 9) What harms are associated with screening for prostate cancer?

Because we found little evidence about key question 9, the harms of screening (one article with inconclusive results), we did not discuss this issue in our article.

Eligibility Criteria for Admissible Evidence

The EPC staff and USPSTF liaisons developed eligibility criteria for selecting relevant evidence to answer the key questions (Appendix Table). We first searched for evidence from RCTs for the efficacy of screening. Because we found no well-conducted and well-analyzed RCT of screening, we then examined case-control and ecologic evidence regarding the overarching key question (key question 1).

For key question 2, concerning the operating characteristics of screening tests, we examined well-conducted systematic reviews and individual studies that started with a primary care or unselected sample without prostate cancer and that compared the findings of one or more screening tests with an adequate reference standard. We also examined evidence of the yield of screening from well-conducted screening programs. For key questions 3 through 6, concerning the effectiveness of various therapies, we required evidence from RCTs. For key questions 7 and 9, concerning the harms of screening or treatment, we required either RCTs or well-controlled studies that included patient reports and the use of a valid measurement instrument. Finally, for key question 8, we searched for evidence of the costs and cost-effectiveness of screening, including models of potential benefits, that considered all appropriate costs and estimates of effectiveness supported by reasonable assumptions based on good evidence.

Literature Search Strategy and Synthesis

The analytic framework and key questions guided our literature searches. We examined the critical literature described in the review published by the USPSTF in 1996 (4) and searched the reference lists of systematic reviews (including Cochrane Library reviews) published since 1993. We then used our eligibility criteria to develop search terms and searched the MEDLINE database for relevant English-language articles that used human subjects and were published between 1 January 1994 and 15 September 2002. We especially looked for articles involving patients whose experience was clearly generalizable to a primary care U.S. population.

The search strategy and results are given in the Appendix Table and in Appendix Figures 2 through 5. All searches started with the term prostate neoplasm and then proceeded by adding further terms as shown in the Appendix Table. The first author reviewed abstracts of all articles found in the searches to determine which met eligibility criteria. Other EPC authors of the full systematic evidence review reviewed all abstracts excluded by the first reviewer. We retrieved the full text of all articles not excluded by both reviewers (Appendix Table).

One reviewer then compared the full text of all retrieved articles against the eligibility criteria and discussed all excluded articles with one of the other reviewers. We included any article that either reviewer judged had met eligibility criteria (Appendix Table). Three of the authors of the systematic evidence review then divided the articles and abstracted data from them, entering the relevant data into predesigned evidence tables (see Appendix B to the systematic evidence review “Screening for Prostate Cancer,” available on the Agency for Healthcare Research and Quality Web site [www.ahrq.gov]). The author who abstracted the articles also graded them using USPSTF criteria (5). The first author read all articles, checked the grading, and discussed crucial ones with a second reviewer. The authors also discussed key articles with the USPSTF liaisons.

Development of the Final Systematic Evidence Review

We presented an initial work plan, including a provisional analytic framework and key questions, to the entire Task Force in September 2000; we also presented interim reports on results of the literature search and the early results of the synthesis of information in December 2000 and March 2001. A draft of the systematic evidence review was submitted for broad-based external peer review in May 2001; the peer review involved experts in the field, representatives of relevant professional organizations, and representatives of organizations and federal agencies that serve as liaisons to the USPSTF. We revised the evidence review as appropriate after receiving peer review comments. The Task Force reviewed all information and voted on a recommendation in June 2001, revising the recommendation and rationale state-
ment in the spring of 2002 after review by involved professional associations and agencies. We then updated the searches, finalized the review, and shortened it for publication.

From University of North Carolina at Chapel Hill, Chapel Hill, and Research Triangle Institute, Research Triangle Park, North Carolina.

Current Author Addresses: Dr. Harris: Cecil G. Sheps Center for Health Services Research, CB# 7590, University of North Carolina School of Medicine, 725 Airport Road, Chapel Hill, NC 27599-7590. Dr. Lohr: Research Triangle Institute, PO Box 12194, 3040 Cornwallis Road, Research Triangle Park, NC 27709-2194.

Appendix Figure 1. Analytic framework for screening for prostate cancer.

Arrows represent steps in the chain of logic connecting screening with reduced morbidity or mortality from prostate cancer. KQ = key question.
### Appendix Table. Eligibility Criteria, Search Strategy, and Results of Searches*

<table>
<thead>
<tr>
<th>Key Question</th>
<th>Inclusion Criteria</th>
<th>Search Terms Used</th>
<th>Articles Identified for Abstract Review</th>
<th>Articles Retained for Full Review</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Efficacy of screening in reducing mortality from prostate cancer</td>
<td>RCT or case–control study or Surveillance (ecologic) study of prostate cancer incidence, morbidity, or mortality over time</td>
<td>“Prostate neoplasms” “Mass screening” “RCT” “Case–control” “Prostate neoplasms” “Incidence” “Mortality” “Trends” “Surveillance”</td>
<td>100 RCT: 1 Case–control: 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Yield of screening tests</td>
<td>Unselected sample without prostate cancer Screening test used for all Result of screening test compared with a valid reference standard</td>
<td>“Prostate neoplasms” “Mass screening” “DRE, PSA” “Diagnosis” “Sensitivity/specificity” “Predictive value” “Reproducibility” “Screening programs”</td>
<td>1905</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>3–6. Health outcomes of treatment</td>
<td>RCT ≤2 years of follow-up ≤75% of patients followed Health outcomes</td>
<td>“Prostate neoplasms” “Therapeutics” “Treatment” “Surgery” “Prostatectomy” “Radiation” “Brachytherapy” “RCT”</td>
<td>656 KQ 3: 1 KQ 4: 0 KQ 5: 2 KQ 6: 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Harms of treatment</td>
<td>Unselected sample with prostate cancer Treated group compared with valid comparison group Randomized trial or adjustment for confounders Not metastatic cancer Valid measures of harms ≥75% of patients followed ≥1 year of follow-up</td>
<td>“Prostate neoplasms” “Therapeutics” “Treatment” “Surgery” “Prostatectomy” “Radiation” “Adverse effects” “Side effects” “Impotence” “Urinary incontinence” “Quality of life”</td>
<td>923</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td>9. Harms of screening</td>
<td>Unselected sample Screened group compared with unscreened group Randomized trial or adjustment for confounders Reliable measure of adverse effects</td>
<td>“Prostate neoplasms” “Mass screening” “Adverse effects” “Anxiety, depression” “Labeling” “Quality of life”</td>
<td>94</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

* DRE = digital rectal examination; KQ = key question; PSA = prostate-specific antigen; RCT = randomized, controlled trial.
Appendix Figure 2. Selection of articles based on key question 1.


Appendix Figure 3. Selection of articles based on key question 2.

Arrows represent steps in the chain of logic connecting screening with reduced morbidity or mortality from prostate cancer.

Appendix Figure 4. Selection of articles based on key questions 3 through 6 (top) and top question 7 (bottom).

Appendix 5. Selection of articles based on key questions 8 (top) and 9 (bottom).

Key question 8 addresses costs and cost-effectiveness of screening, and key question 9 deals with the harms of screening.